

ある電子図書館の運用と統計

今井正和 新麗, 羽田久一, 西村亨, 砂原秀樹, 千原國宏
奈良先端科学技術大学院大学
〒 630-0101 奈良県生駒市高山町 8916-5
imai@is.aist-nara.ac.jp

概要

本報告では、平成8年4月より運用を開始した奈良先端科学技術大学院大学附属電子図書館について、そのアクセス記録をもとにした利用状況の報告を行う。アクセス記録によると、開館から半年ほどの間は日本国内からのアクセスが顕著であった。それ以降は学内からのアクセスの方が多くなり、また増加傾向が続いている。さらに閲覧された電子資料を見てみると、学内生産物の閲覧が比較的多く、大学附属図書館の重要な機能の一つである大学からの情報発信に大きく寄与していることが観察できた。従来からその重要性については議論されてきたところではあるが、電子図書館とすることにより実現できた。

キーワード

電子図書館, 運用, 利用統計

Operation and Report of A Digital Library

Masakazu Imai, Rei Atarashi, Hisakazu Hada, Toru Nishimura,
Hideki Sunahara, and Kunihiro CHIHARA
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara, 630-0101 Japan
imai@is.aist-nara.ac.jp

Abstract

This paper reports statistics of Digital Library of Nara Institute of Science and Technology. NAIST Digital Library records its access from its opening. The access log shows that more than 50% of access was from Japanese domains which does not include NAIST during first half year. After the first half year, local access is more than that of domestic access and increasing constantly. Documents written in NAIST are well accessed documents. An important function of an university library is distributing documents which are written in the university. The access log significantly shows that our digital library did well the function.

Keywords

Digital Library, Operation, Statistics

1. はじめに

現在、世界各国では電子図書館をテーマとしたプロジェクトが多数行われている。米国では1994年より続けられていたNSF/NASA/DARPA支援のDigital Library Initiative (DLI) プロジェクトも1998年8月で終了し、DLI2が始まろうとしている。DLIプロジェクト以外にもLibrary of Congressなどが支援して多数のプロジェクトが行われている。これらのプロジェクトではさまざまな文献を対象として電子図書館が実験的あるいは本格的に構築されている。また、国内では学術情報センタが学会論文誌を中心に電子図書館サービスを始め、京都大学や筑波大学などでも貴重書を中心とした電子図書館プロジェクトが進んでいる。1991年に開学した奈良先端科学技術大学院大学では、開学当初よりその附属図書館を電子図書館として構築することが検討されていた。幸いにも、関係各方面の協力を得て平成7年度には構築のための経費が認められ、平成8年4月より日本で初めての大学附属電子図書館として運用を開始した。本報告では、この奈良先端科学技術大学院大学附属電子図書館の運用開始からの2年間の運用状況を報告する。

2. 奈良先端科学技術大学院大学附属電子図書館

本学の電子図書館についてはその概要を文献[1]に述べているが、文献[1]での報告後に若干の改良がなされている。このため、本章では改良された点を含めて簡単にその概要を述べる。また、文献[2]にはより詳細な記述がある。

2. 1 電子図書館の目的

奈良先端科学技術大学院大学は先端科学技術の研究を行うため、3つの研究科（情報科学研究科、バイオサイエンス研究科、物質創成科学研究科）を中心に構成されている。その附属図書館は学内の研究科において行われる研究活動のサポートが最も重要な目的になっている。各研究科では最先端の研究が行われているため、図書よりも学術雑誌、各種データベースが重要な情報源になっている。さらに、これらの情報源についても最新の情報が要求される。また、学内で生産された論文などの学術情報を学外に発信することも重要な目的の一つである。電子図書館を構築するにあたり、この本学附属図書館が果たすべき使命を電子的に実現することを最も重要な目標とした。

2. 2 システム構成

電子図書館システムは一次情報入力、蓄積、検索、AVの主要なサブシステムから構成されている。また、利用者は手元の計算機からWorld Wide Webのブラウザ（以下、単にブラウザとする）を用いて電子図書館を利用する。以下、それぞれのサブシステムについて簡単に述べる。

2. 2. 1 一次情報入力サブシステム

一次情報入力サブシステムは、各種メディアから一次情報を電子図書館に入力するためのシステムである。紙メディア、電子メディアからの入力に対応しており、Plain Text, HTML, Postscript, PDFの各電子メディアの形式に対応している。分当たり25ページの電子化が可能である。

紙メディアの入力は、自動ページ送り機構を備えたスキャナにより行う。一次情報を画像情報として入力したあと、OCRソフトウェアを用いて画像情報から文字情報を抽出している。画像入力は400DPIの解像度で行われ、OCRと印刷用データとして用いる。また、このデータの解像度を100DPIに変換して閲覧用データとしている。スキャンに際しては、基本的には自動ページ送り機構を備えたコピー機で一旦コ

ピーを作成し、閉じられている紙を個別のページに変換する。紙質などの問題からコピー機でコピーが作成できないものについては、雑誌の背を裁断し、個別の紙としてスキャンする。

電子メディアの入力にはその供給形態によって2つの方法がある。電子メディアを直接FTPなどにより得ることができる場合は、作業領域にデータを格納し、データ形式に応じた処理を行う。CD-ROMによって供給される場合には、読み取り装置を用いてデータを読み出し、作業領域に格納し、データ形式に応じた処理を行い、システムに登録する。

入力される文献の二次情報は、別途入力される。このとき、電子形態で利用できるものがあれば、それを利用し作業量の低減をはかっている。

2. 2. 2 一次情報蓄積サブシステム

一次情報蓄積サブシステムでは、電子図書館に入力された電子形態での一次情報とAV情報を蓄積するための記憶システムを提供する。1998年10月現在、記憶システムの容量は書籍の一次情報用に4.9TB、AV情報用に180GBである。

書籍の一次情報を格納する記憶システムは、利用者からのアクセス要求に迅速に反応するためにはすべてをアクセス速度の速い磁気ディスクで構成するのが利用的である。しかし、その場合はコスト面で大きな問題となる。一方、磁気テープはアクセス速度は遅いが、ビット当たりの記憶コストが磁気ディスクに比べて非常に小さい。そのため、本学の電子図書館では、磁気ディスクと磁気テープを併用したマイグレーションシステムを用いて、コストと実用性を折衷させたものである。

AV情報は、通常のFDDIインタフェースを備えた記憶システムに蓄積する。この記憶システムはFDDIネットワークインタフェースを用いた通常のファイルサーバであり、同時に3チャンネル以上のビデオ再生要求に耐えることができる。

2. 2. 3 検索サブシステム

検索サブシステムは、利用者とのインタフェースをとるためのHTTPサーバも兼ねている。一次情報入力システムにより入力された図書情報は本サブシステムにあるデータベースソフトウェアによりデータベース化される。通常の図書館が備えるOPAC (Open Public Access Catalog) の機能に加え、電子化された書籍に対する全文検索機能や、その他の主要な電子図書館機能を実現している。また、利用者からの閲覧要求を処理し、それが適切なものであるかの認証を行い、認証された要求に対しては閲覧のための一次情報を一次情報蓄積サブシステムから読み出し、クライアントに送出する。

2. 2. 4 AVサブシステム

重要な研究成果の表現法としてビデオがある。奈良先端科学技術大学院大学附属電子図書館では、このようなビデオ情報 (AV情報) も重要な研究を支援する情報であると考え、電子化し、利用者に提供している。AV情報の作成を支援するため、テレビカメラ、ビデオ機器、ノンリニアビデオ編集機などがある。さらに、システムではないが、簡単なスタジオ設備もある。作成されたAV情報を電子化するため、MPEG-2方式によるビデオエンコーダがある。標準では8Mbpsの量子化速度でAV情報を電子化する。エンコーダは12Mbpsまでの量子化速度に対応しており、電子化するAV情報の特徴にあわせて量子化速度を選択している。

3. 電子図書館の運用統計

本学附属電子図書館システムでは、サーバソフトウェアから様々なログ情報が記録されている。それらのアクセス記録から、今回はHTMLサーバのログ情報をなどをもとに簡単な集計を行った。本章ではそれらの集計結果を示す。

3.1 書籍の電子化状況

まず最初に電子図書館における資料の電子化状況について述べる。電子図書館運用開始後、四半期ごとの書籍の電子化状況を表1、図1に示す。一次情報入力サブシステムの処理能力は1分間に25ページ電子化できる能力を持つが、以下のような理由により現在のところ処理能力が定常的に飽和するにはいたっていない。

表1 資料の電子化

	(四半期ごとの累積)									
	'96.4/1	'96.4/2	'96.4/3	'96.4/4	'97.4/1	'97.4/2	'97.4/3	'97.4/4	'98.4/1	'98.4/2
雑誌冊子体	11,522	27,017	40,836	67,493	94,220	118,661	135,702	159,749	181,210	201,817
雑誌CD-ROM	0	1,222	20,745	91,638	95,626	95,751	101,793	121,685	211,465	221,380
図書	0	0	0	507	1,948	3,185	3,354	3,354	4,483	10,122
学内生産物	1,062	1,062	1,062	1,062	3,519	3,624	9,568	9,568	9,922	14,883

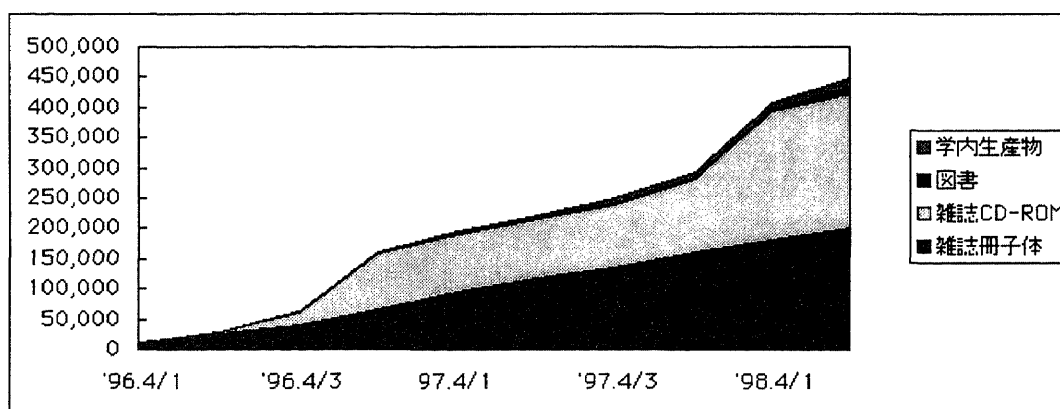


図1 四半期ごとの電子化資料の数の推移

- 著作権許諾の問題
- データバックアップ速度

平成8年の開館以来積極的に資料の電子化のための著作権交渉を行ってきた。その結果、これまでに雑誌135タイトル以上、図書、学内生産物あわせて340タイトル以上の電子化許諾を得ている。今年度から物質創成科学研究科が学生受け入れを開始したため、今後許諾を得る雑誌のタイトル数は増加すると見込まれる。また、これまでは条件が折り合わないため電子化を行っていない雑誌もかなりあり、今後条件が整ってくれば電子化すべき資料の数は増大の一步をたどると考えられ、現有の処理能力は過大なものであるとは考えていない。また、本学附属電子図書館では、電子化された資料を一次情報蓄積サブシステムに

格納する際に、バックアップ作成を考えテープシステムに書き込む操作を同時に行っている。このバックアップ作成のためにテープにデータの書き込みがあるが、この速度が一次情報入力サブシステムの電子化速度よりも遅いため、電子化の速度には限界がある。ただし、定期的に資料の電子化作業が行われている場合にはテープでのデータ書き出しの速度は十分に速く、これによって電子化作業が遅延する事態は発生していない。なんらかの原因で入力作業が長期にわたって停止した場合、電子化作業が定常状態に復帰する場合に多少の時間がかかってしまう問題がある。

3. 2 アクセス状況

開館以来の本学附属電子図書館へのアクセス状況を報告する。まず、開館以来のアクセス数の変化を図2に示す。これはHTTPサーバ(Appatch)が残すログファイルからヒット数をカウントしたものである。図3にはクライアントの所属ドメインで分類したヒット数の変化をグラフ化したものである。図2から平成8年11月頃から平成10年3月頃までは1ヶ月に10万件弱で推移し、平成10年4月以降は12万件前後で推移していることが読み取れる。開館から平成8年10月頃までは、開館を積極的にマスメディアなどで広報したためアクセス件数が多いが、それ以降は定常状態に入りつつあると考えられる。平成10年4月よりアクセス件数の急増が見られるのは、物質創成科学研究科が学生の受け入れを開始した時期と符合している。このため、アクセス件数は来年度も増加する傾向が続くと考えられる。図3に示した所属ドメインごとによるアクセス件数の変化を見るとこの傾向がより顕著に観察される。このグラフから、開館後の半年間のアクセス件数は学外(日本国内)からのアクセスが多かった。これに対して、学内からのアクセスは増減をくり返しながらも全体としては順調に増加を続けていると考えられる。また、利用者が予め登録したキーワードを用いて、図書館システムに入力が完了したデータを1日に1回夜半に検索し、その結果を新着資料として電子メールで通知する機能がある。この機能の利用が進んできている可能性も否定できない。平成10年8月に国内からのアクセスが急増している時期があるが、これは本学の自己評価報告書が電子図書館で公開され、電子図書館を通じて閲覧されたためである。

海外からのアクセスは低調である。本学附属電子図書館は、その性格上日本語と英語の両方の文献を保有している。英語で記述されたページが用意され、海外からのアクセスに対応することがはかられているが、検索結果には当然ながら日本語が含まれることが多くなる。しかし、海外のクライアントで日本語フォントを持つものはそれほど存在するとは思えないし、また日本語の文字コードが送られてきた場合に正しく対処できるようにブラウザの設定をしていることは稀であると考えられる。その結果、海外におけるクライアントでは送られてくる日本語文字コードのため、ブラウザが面が著しく乱され、継続した使用がなされないものと思われる。今後、日本語文字コードが混入しないようなページ設計が必要である。

アクセスの季節変動を見るため、図4に各年度ごとのアクセスの変化をグラフ化したものを示す。これによると、6月頃、11月頃、2月頃にピークが見られる。これは次のように解釈できる。4月に研究室へ配属された学生が電子図書館に対する興味からアクセスを行うことにより6月頃に一つのピークが形成される。夏休みの時期にはアクセスが減るが、秋には研究の動向調査が行われ、ふたたびピークが形成される。さらに、年明けには修士論文作成のため文献調査が行われる。

3. 3 資料の被アクセス動向

本学電子図書館では、どのような資料がアクセスされたかを示すログを残している。図5と図6はそれぞれ、雑誌、図書(学内生産物を含む)の月ごとのアクセス数の変動を示す。

図5に示した雑誌のアクセス数のグラフから、順調にアクセス数が増加していることが読み取れる。平成9年11月にアクセス数が急増しているのは、事故があったため、データの整合性を集中的に検査したため

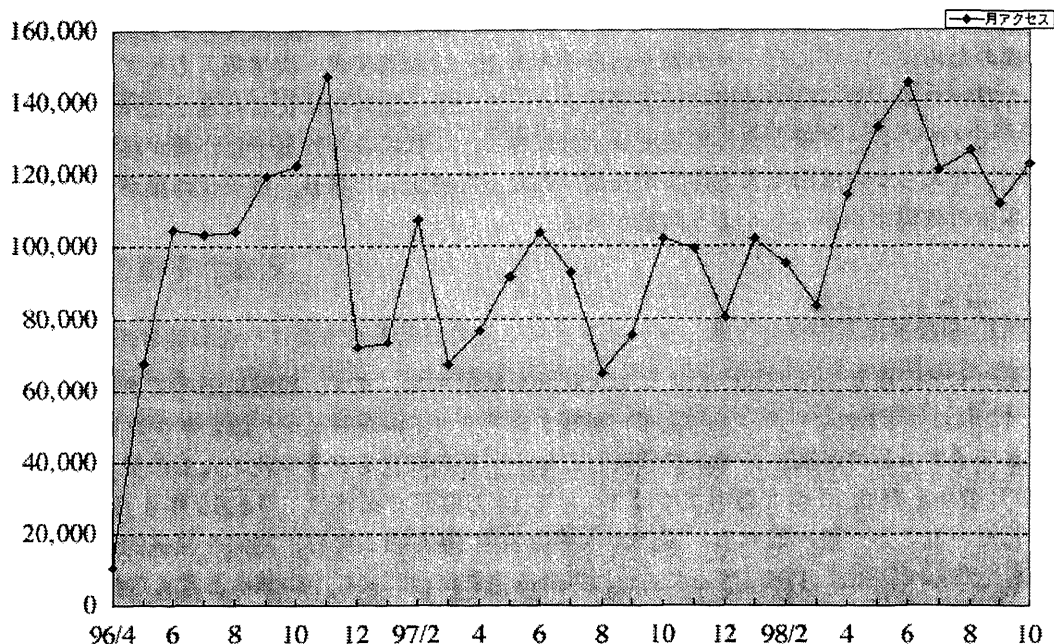


図2 電子図書館のアクセス件数の推移

である。図6に示した図書（学内生産物を含む）のアクセス数では，単調増加の傾向が読み取れない。平成10年1月以降は，本学電子図書館についての大学が作成したレポート「NAIST 電子図書館レポート'97」が電子化され，一般に公開したことによるアクセスが増加している。また，平成10年6月以降の急激な増加は，本学の自己評価報告書「新世紀へ向けて：NAISTの検証」が電子化され，一般に公開されたことによる。また，著作権許諾の関係から，現実には図書としては修了した学生の修士論文が多数を占めており，図6でのアクセス数のほとんどはこれら修士論文の閲覧である。

本学電子図書館で電子化されている資料のうち，どのようなものがよく閲覧されているかを示したものが表2と表3である。表2は雑誌のうち，アクセス数が多いもの上位10点をその総アクセス数とともに示す。表2から，雑誌がカバーする範囲が広いもののアクセスが多いことがわかる。また，表3からは，本学が公開した自己評価報告書や辞書的なもののアクセスが多いことがわかる。なお，表3で第4位から第10位までは本学修了生の修士論文である。これから電子図書館が大学が生産する情報の発信に大きな貢献をしていることが読み取れ，これまでの図書館では問題となっていた情報発信機能を実現しているといえる。

4. まとめ

本報告では，奈良先端科学技術大学院大学附属電子図書館が平成8年4月に開館してからのアクセス動向について述べた。電子図書館を運用しているところは世界的に見ても少なく，運用することによって始めて明らかになる点もいくつかあった。今回の電子図書館によって，

- 記憶システムの構成法
- 英文ページへの日本語文字コードの混入

といった問題点が判明した。いずれも，良く考えてみると当たり前のようなことであるが，見落とされがちなことであると考える。次期バージョンではこれらの問題点に対処する予定である。

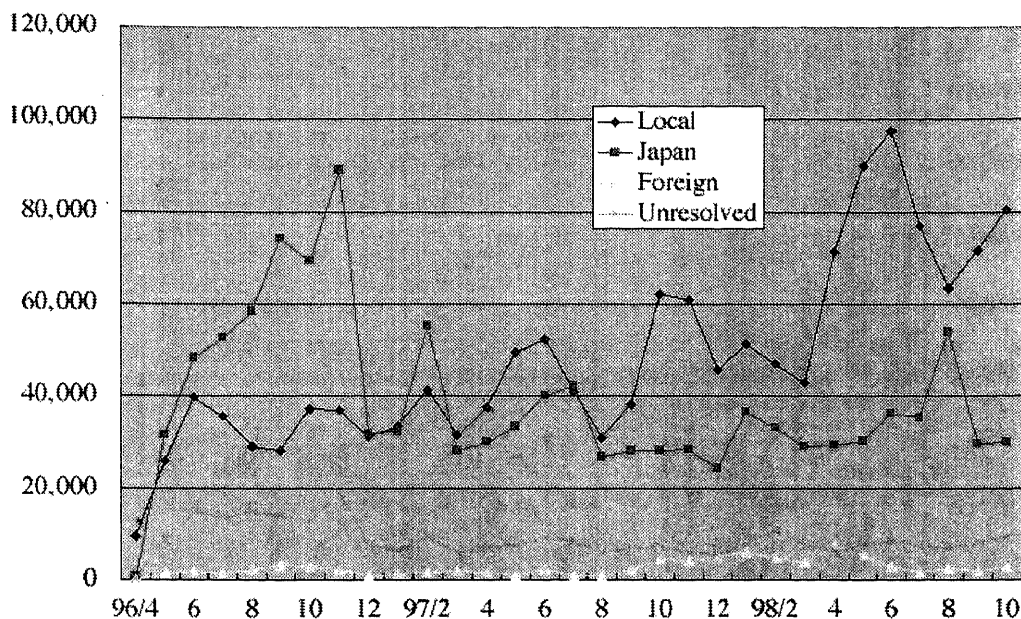


図3 クライアントの所属ドメインごとのアクセス件数

機能面では、これまでの図書館では難しかったと思われる情報発信機能が有効に機能していることが示された。大学図書館は学内で行われる研究に必要な情報の提供のみならず、電子図書館とすることによってその情報発信機能が強化される。今後、学内生産物を積極的に電子化し、図書館から公開することがますます重要になってくるとと思われる。

学内の利用者からは一部、「使用法が分からない」や「使いづらい」といった種類の意見も聞かれる。これらの意見はユーザインタフェースに関わる部分であるが、我々は現在のものが最上のものであるとは考えていない。現在のシステムは、平成8年度から運用が開始されているため、平成7年の時点の技術で一般競争入札に馴染む形で構築されている。今後も一般競争入札の制約はあるが、技術の進展を積極的に取り入れたシステム構築を続けていきたい。電子図書館は現在研究テーマとしても最も新しいものでもあり、ユーザインタフェースはその中でも重要なものである。幸い、平成10年8月より本学附属図書館には学内処置で研究開発室が設置され、専任教官が配置された。電子図書館に関連する研究を行う環境が整備されつつある。研究開発室では電子図書館に関連した研究を行い、その成果を将来の電子図書館システムに反映することをめざしている。

参考文献

- [1] 今井正和：“大学における電子図書館の構築 —奈良先端科学技術大学院大学電子図書館—”，第8回デジタル図書館ワークショップ，1996年10月
- [2] 奈良先端科学技術大学院大学附属電子図書館編：「NAIST 電子図書館レポート'97」，平成9年，<http://dlw3.aist-nara.ac.jp/>

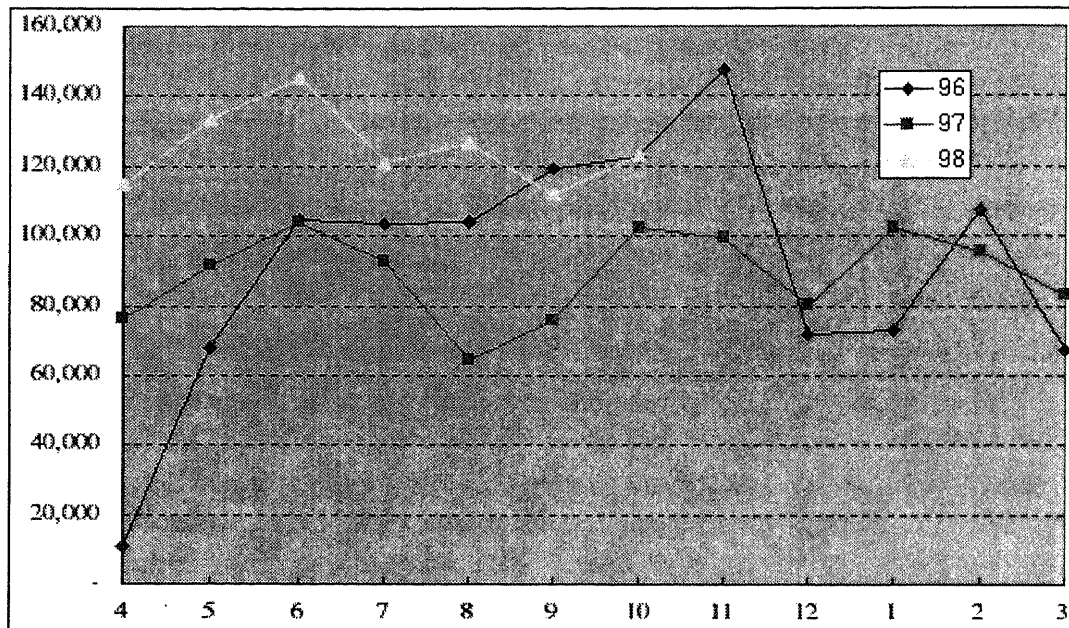


図4 季節によるアクセス件数の変動

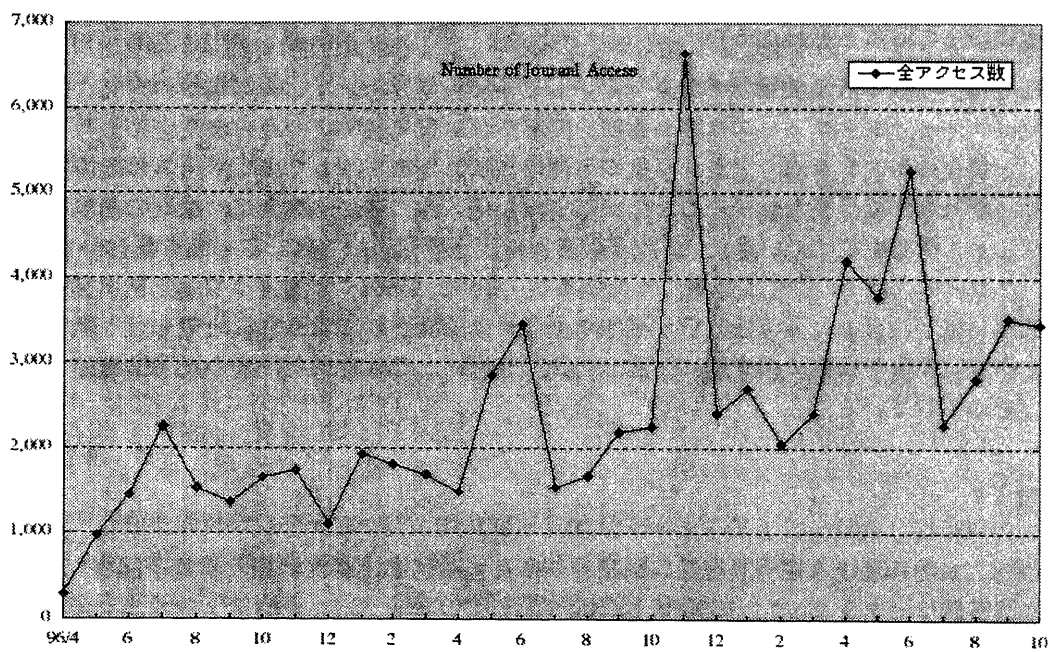


図5 雑誌のアクセスの変動

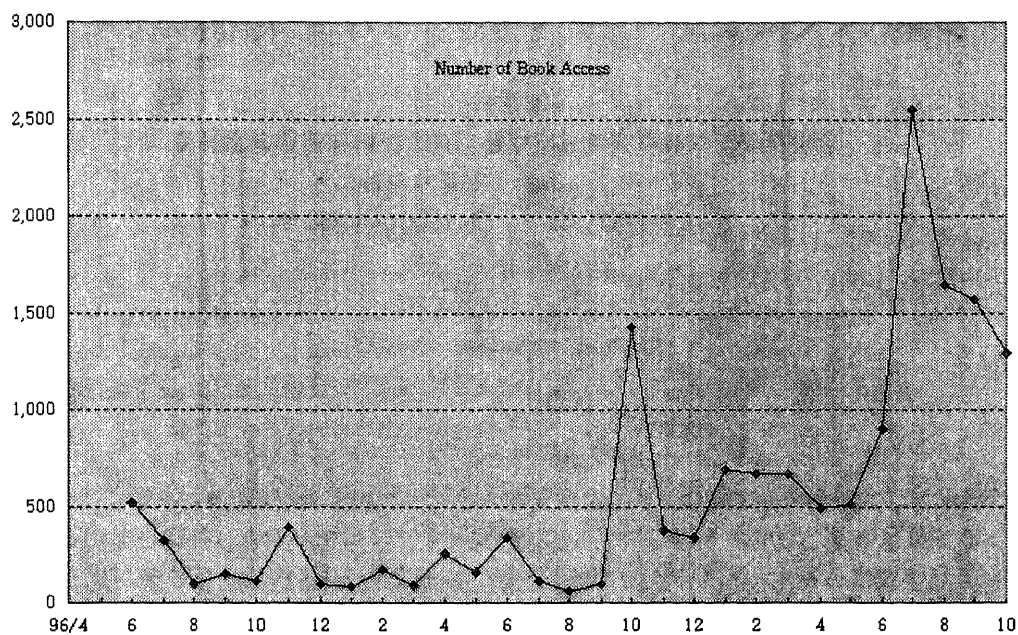


図6 図書（学内生産物を含む）のアクセス変動

表2 よくアクセスされる雑誌

	アクセス数
日経コンピュータ	4,590
日経エレクトロニクス	4,441
情報処理学会論文誌	4,199
電子情報通信学会論文誌. D-2, 情報・システム 2-情報処理	3,009
日経パソコン	2,970
日経バイト	2,635
日経ビジネス	2,292
日経コミュニケーション	2,086
Plant molecular biology	2,074
日経CG	2,004

表3 よくアクセスされる図書

	アクセス数
新世紀へ向けて : NAISTの検証	5023
NAIST電子図書館レポート'97	2711
「NTT技術ジャーナル」にみる最新情報通信用語集	429
コンピュータの設計とテスト	406
専門文書の概念構造の抽出と知識メディア化	348
Studies on merit functions for the variational inequality problem and nonlinear complementarity problem	311
オブジェクトモデルの変換機能をもつ情報検索方式	299
ATM制御ソフトウェアの再利用における機能仕様の不整合検出方法	297
ネットワーク環境におけるテレカンファレンスの音場制御について	279
共有仮想空間の知的利用のための定性的システム同定法と2レベルオントロジーに関する研究	275